TESTING PROGRAM COMPONENTS USING THE MULTIPHASE OPTIMIZATION STRATEGY (MOST)

Linda M. Collins, Ph.D.
The Methodology Center and Department of Human Development & Family Studies
Penn State

Presented at
What Works, Under What Circumstances, and How?
Methods for Unpacking the “Black Box” of Programs and Policies
Office of Planning, Research and Evaluation
Administration for Children and Families
Washington, DC
September 3, 2014
OUTLINE

• What’s wrong with business as usual?
• What is MOST? What is optimization?
• Selecting an approach to examine individual components
• Take-home messages
HOW BEHAVIORAL INTERVENTIONS ARE TYPICALLY DEVELOPED AND EVALUATED

• Intervention components are chosen based on scientific theory, clinical experience, etc.

• Combined into a package

• Package is evaluated via a randomized controlled trial (RCT)

• Let’s call this the treatment package approach
TREATMENT PACKAGE APPROACH

Behavioral intervention

Evaluation via RCT
What’s wrong with evaluating a treatment package via an RCT?

Absolutely nothing!
THE RCT IS BEST SUITED FOR

• Determining whether a treatment package performs better than
  – A control or comparison group
  – An alternative intervention
TREATMENT PACKAGE APPROACH

Behavioral intervention

Evaluation via RCT
WHAT THE RCT CANNOT NOT TELL US

An RCT that finds a significant effect DOES NOT tell us

• Which components are making positive contributions to overall effect

• Whether the inclusion of one component has an impact on the effect of another

• Whether a component’s contribution offsets its cost

• Whether all the components are really needed

• How to make the intervention more effective, efficient, and scalable
WHAT THE RCT CANNOT NOT TELL US

An RCT that finds a non-significant effect DOES NOT tell us

- Whether any components are worth retaining
- Whether one component had a negative effect that offset the positive effect of others
- Specifically what went wrong and how to do it better the next time

What is the alternative to sole reliance on the RCT?
OUTLINE

• What’s wrong with business as usual?
• What is MOST? What is optimization?
• Selecting an approach to examine individual components
• Take-home messages
AN ALTERNATIVE: THE MULTIPHASE OPTIMIZATION STRATEGY (MOST)

• An integration of ideas from engineering and the behavioral sciences

• A framework for development, optimization, and evaluation of behavioral interventions

• MOST INCLUDES THE RCT along with other types of experimentation

• Using MOST it is possible to engineer an intervention to meet a specific criterion
MULTIPHASE OPTIMIZATION STRATEGY (MOST)

Empirically based optimization

Optimized behavioral intervention

Evaluation via RCT
MULTIPHASE OPTIMIZATION STRATEGY (MOST)

Empirically based optimization

Component

Optimized behavioral intervention

Evaluation via RCT
MULTIPHASE OPTIMIZATION STRATEGY (MOST)

Empirically based optimization

Optimized behavioral intervention

Evaluation via RCT
Evaluation and optimization: Both important; not the same thing.

<table>
<thead>
<tr>
<th>Optimization: Is the intervention the <strong>best possible, given constraints?</strong></th>
<th>Evaluation: Is the intervention’s effect <strong>statistically significant?</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>No</td>
<td>May wish to optimize to improve effect size</td>
</tr>
<tr>
<td>Yes</td>
<td>Intervention can probably be improved</td>
</tr>
<tr>
<td>Yes</td>
<td>Different intervention strategy needed</td>
</tr>
<tr>
<td>Yes</td>
<td>What we should be aiming for</td>
</tr>
</tbody>
</table>
OUTLINE

• What’s wrong with business as usual?
• What is MOST? What is optimization?
• Selecting an approach to examine individual components
• Take-home messages
MULTIPHASE OPTIMIZATION STRATEGY (MOST)

Preparation

Optimization

Data for future secondary analyses

Expect optimized intervention to be effective?

Evaluation

Optimized intervention effective?

Release optimized intervention
EXAMPLE APPLICATION OF MOST

- *HealthWise* school-based ATOD/HIV prevention intervention developed by Drs. Linda Caldwell and Ed Smith at Penn State
- Has previously been evaluated in South Africa
- Metropolitan South Education District in South Africa wants to implement HealthWise in all its schools
- Question: how to maintain fidelity?
- Metro South allowed us to conduct an experiment
COMPONENTS

• Enhanced teacher training
  – Standard training (one and one-half days) vs. enhanced (three days + two additional days four months later)

• Structure, support, and supervision
  – No additional vs. additional (e.g., weekly text messages; monthly visits from support staff; option to call support staff with questions as needed)

• Enhanced school climate
  – No climate enhancement vs. climate enhancement (e.g., form committee of parents and teachers to promote HealthWise; develop visuals; issue newsletter)
SELECTING AN EXPERIMENTAL APPROACH FOR THE OPTIMIZATION PHASE OF MOST

• We decided to conduct a factorial experiment. Why?
• Enables examination of individual component effects AND
• Statistical power achieved with smaller sample sizes than alternative designs
• BUT they also usually require more experimental conditions than we may be accustomed to
• Experiment uses all 56 schools in district
SELECTING AN EXPERIMENTAL APPROACH FOR THE OPTIMIZATION PHASE OF MOST

- **NOTE:** ANY experimental design is a possibility
- The SMART, a special case of the factorial experiment, can be used
- Or other experimental designs besides factorial designs
- Or other approaches such as dynamical systems ID or machine learning
COMPARISON OF RESOURCE REQUIREMENTS FOR POWER $\geq 0.8$, MEDIUM EFFECT SIZE

<table>
<thead>
<tr>
<th>DESIGN</th>
<th>N</th>
<th># experimental conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 individual experiments</td>
<td>168</td>
<td>6</td>
</tr>
<tr>
<td>Comparative treatment experiment</td>
<td>112</td>
<td>4</td>
</tr>
<tr>
<td>Factorial experiment</td>
<td>56</td>
<td>8</td>
</tr>
</tbody>
</table>
HealthWise component selection experiment in South African school district. 56 schools in all; 7 schools assigned to each experimental condition

<table>
<thead>
<tr>
<th>Experimental condition</th>
<th>N of schools</th>
<th>HealthWise program</th>
<th>Training</th>
<th>Structure, support, &amp; supervision</th>
<th>Enhanced school climate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>7</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>7</td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>3</td>
<td>7</td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>4</td>
<td>7</td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>5</td>
<td>7</td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>6</td>
<td>7</td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>8</td>
<td>7</td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
</tbody>
</table>
HOW CAN 7 SCHOOLS PER EXPERIMENTAL CONDITION BE ENOUGH?

• In a factorial experiment you DO NOT compare individual conditions
• Each main effect estimate is based on ALL schools
• Power driven by OVERALL $N$, NOT per-condition $n$
**Main effect of Training** is mean of (5,6,7,8) vs. mean of (1,2,3,4). Note that all 56 schools are used in estimating the main effect.

<table>
<thead>
<tr>
<th>Experimental condition</th>
<th>N of schools</th>
<th>HealthWise program</th>
<th>Training</th>
<th>Structure, support, &amp; supervision</th>
<th>Enhanced school climate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>7</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>7</td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>3</td>
<td>7</td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>4</td>
<td>7</td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>5</td>
<td>7</td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>6</td>
<td>7</td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>8</td>
<td>7</td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
</tbody>
</table>
Main effect of Structure, support, & supervision is mean of (3,4,7,8) vs. mean of (1,2,5,6).

Note that all 56 schools are used in estimating the main effect.

<table>
<thead>
<tr>
<th>Experimental condition</th>
<th>N of schools</th>
<th>HealthWise program</th>
<th>Training</th>
<th>Structure, support, &amp; supervision</th>
<th>Enhanced school climate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>7</td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>2</td>
<td>7</td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>3</td>
<td>7</td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>4</td>
<td>7</td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>5</td>
<td>7</td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>6</td>
<td>7</td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>8</td>
<td>7</td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
</tbody>
</table>
Main effect of Enhanced school climate is mean of (2.4.6.8) vs. mean of (1,3,5,7).

Note that all 56 schools are used in estimating the main effect.

<table>
<thead>
<tr>
<th>Experimental condition</th>
<th>N of schools</th>
<th>HealthWise program</th>
<th>Training</th>
<th>Structure, support, &amp; supervision</th>
<th>Enhanced school climate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>7</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>7</td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>3</td>
<td>7</td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>4</td>
<td>7</td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>5</td>
<td>7</td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>6</td>
<td>7</td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>8</td>
<td>7</td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
</tbody>
</table>
ENGINEERING THE INTERVENTION

• Experiment will give us
  – Main effect of each individual intervention component on outcome variables
  – Also interactions between intervention components

• This information will be used to select the best set of the three components
OUTLINE

• What’s wrong with business as usual?
• What is MOST? What is optimization?
• Selecting an approach to examine individual components
• Take-home messages
TAKE-HOME MESSAGES

• It is possible and feasible to conduct experimentation to examine individual intervention components

• Possible even when cluster randomization is necessary

• Factorial experiments can be really efficient when properly applied and analyzed
FOR MORE INFORMATION:
http://methodology.psu.edu/ra/most

This website has
– Suggested reading
– FAQ
– Advice for people writing grant proposals involving MOST

For a brief conceptual introduction, go to YouTube and see my 12-min TED talk: Reducing smoking deaths: Is it rocket science?
A FEW ARTICLES


